

(11)Publication number : 11-259643
(43)Date of publication of application : 24.09.1999

(51)Int.Cl. G06T 1/00
G06F 15/18
// H04N 7/18

(21)Application number :	10-269987	(71)Applicant :	MITSUBISHI ELECTRIC INF TECHNOLOG CENTER AMERICA INC
(22)Date of filing :	24.09.1998	(72)Inventor :	BRAND MATTHEW E

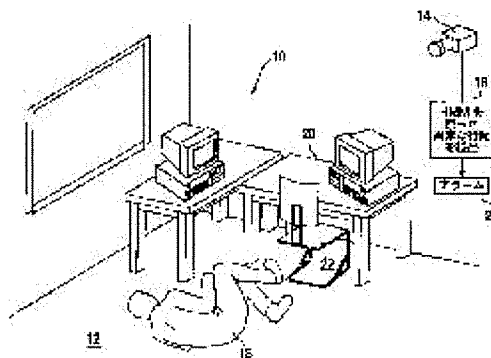
(30)Priority
Priority number : 98 16375 Priority date : 30.01.1998 Priority country : US

(54) ACTION DETECTION SYSTEM

(57)Abstract:

PROBLEM TO BE SOLVED: To automatically acquire probability models for standard actions through the use of hidden Markov models by observing the world through a video camera and other sensors, and to detect an abnormal action by inferring the kinds of activities which are currently monitored by using the models.

SOLUTION: In an embodiment, a system consists of image analysis, model acquisition, and signal analysis as three constituent parts. In the image analysis, respective frames of video are deformed into vectors of a sufficient number to represent the operation of an object 18 in front of the camera 14 and a series of vectors establishes one 'signal' for each frame of the video. In the model acquisition, signals are analyzed so as to obtain parameters as to dynamic probability models of the scene in front of the camera 14. In the signal analysis, the models are used to classify and/or detect abnormality of signals generated speedily by image analysis of new video. Consequently, the processing time can greatly be shortened.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

特開平11-259643

(43) 公開日 平成11年(1999) 9月24日

(51) Int. Cl. ⁶ 識別記号

G06T 1/00

G06F 15/18 560

// H04N 7/18

F I

G06F 15/62 380

15/18 560 Z

H04N 7/18 D

審査請求 未請求 請求項の数 8 O L (全13頁)

(21) 出願番号 特願平10-269987

(22) 出願日 平成10年(1998) 9月24日

(31) 優先権主張番号 09/016375

(32) 優先日 1998年1月30日

(33) 優先権主張国 米国 (US)

(71) 出願人 597067574

ミツビシ・エレクトリック・インフォメイ
ション・テクノロジー・センター・アメリ
カ・インコーポレイテッドMITSUBISHI ELECTRIC
INFORMATION TECHNO
LOGY CENTER AMERICA
, INC.アメリカ合衆国、マサチューセッツ州、ケ
ンブリッジ、ブロードウェイ 201

(74) 代理人 弁理士 曾我 道照 (外6名)

最終頁に続く

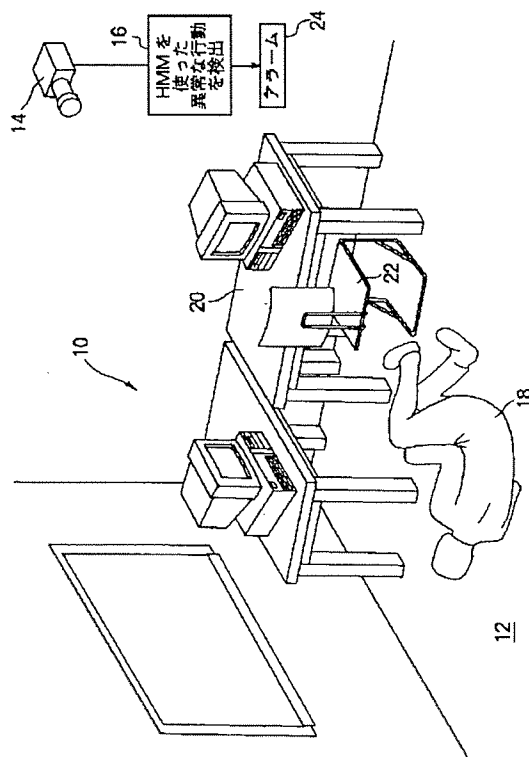
(54) 【発明の名称】 行動検出システム

(57) 【要約】

【課題】 ビデオカメラ、他のセンサを介して世界を観察し、隠れマルコフモデルの使用を通して標準的行動の確率モデルを自動的に習得し、そのモデルを使って、現在監視下にある活動の種類を推論して異常な行動を検出することを課題とする。

【解決手段】 一実施例においては、システムは3つの構成部分である画像分析、モデル習得および信号分析からなる。画像分析においては、映像の各フレームは、カメラの前の対象物の動作を描写する数のベクトルに変形され、一連のかかるベクトルは、映像の各フレームについて1個の「信号」を確立する。モデル習得においては、信号は、カメラの前の場面の力学の確率モデルについてパラメータを得るために分析される。信号分析においては、モデルは、新たな映像の画像分析によって迅速に作り出される信号における異常を分類および/または検出するために使用される。

【効果】 処理時間を大幅に削減できる。



【特許請求の範囲】

【請求項 1】 入力信号から標準的行動の確率モデルを自動的に習得する行動検出システムにおいて、隠れマルコフモデルを提供する手段と、前記入力信号とエントロピー的プライア(entropy prior)に基づいて前記隠れマルコフモデルをトレーニングすることにより、前記確率モデルを生成する手段とを備えた行動検出システム。

【請求項 2】 前記エントロピー的プライア(entropy prior)は θ^A の形式であり、したがって殆ど情報として役立たない内容を有する前記隠れマルコフモデルからのパラメータを消去する請求項 1 記載の行動検出システム。

【請求項 3】 前記トレーニングは、前記隠れマルコフモデルに推定されたパラメータを提供する手段を含み、前記推定されたパラメータを提供する手段は、エントロピー的プライア(entropy prior)を含む請求項 2 記載の行動検出システム。

【請求項 4】 前記隠れマルコフモデルをトレーニングする手段は、監視下にある場面对応した映像のフレームを提供するビデオカメラと、前記映像の各フレームを前記ビデオカメラの前の対象物の動作を説明する数字のベクトルに変形させる手段とを含み、一連の前記ベクトルは前記入力信号を備える請求項 1 記載の行動検出システム。

【請求項 5】 前記トレーニング手段は、前記確率モデルについてのパラメータを得るための前記エントロピー的プライア(entropy prior)を含む請求項 4 記載の行動検出システム。

【請求項 6】 前記エントロピー的プライア(entropy prior)は θ^A の形式である請求項 5 記載の行動検出システム。

【請求項 7】 標準的行動に関する前記入力信号を分類する前記隠れマルコフモデルを含む手段をさらに含む請求項 5 記載の行動検出システム。

【請求項 8】 前記分類手段によって確立された標準的行動に基づいて、前記入力信号から異常な行動を検出する、前記分類手段に接続される手段をさらに含む請求項 7 記載の行動検出システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明は、ビデオカメラによって検出される個人の標準的行動の確実な検出に関し、特にエントロピー的プライア(entropy prior)でトレーニングされた隠れマルコフモデル(Hidden Markov Model)を使って、場面の実時間の映像分析を用いて人間の異常な行動を検出する行動検出システムに関するものである。

【0002】

【従来の技術】過去には、時間と行動の研究、セキュリティや一般的な監視の目的のために、個人の動きを検出して分類することに相当な関心があった。かかる監視システムの多くは、予め選択されているか予め定められた行動を検出することができるように、演繹的知識を必要としている。これらのシステムにおいては、ビデオカメラに捉えられた動作は、ヒストグラムまたは他の種類のテンプレートと比べられる。しかし、これらのシステムは全て、ヒストグラムまたはテンプレートを作成するために検出されることを求められる標準的行動および非標準的行動という観念を要求する。

【0003】その結果、隠れマルコフモデル(Hidden Markov Model)の使用を通して、視覚から活動の標準的モデルを習得することに大きな関心がある。しかし、これらのモデルを使用することは、極端に精度が低いために不満足な結果に帰結する。

【0004】注目すべきは、「画像のシーケンスからの意味論的領域の生成(Generation of Semantic Regions from Image Sequences)」というタイトルの論文において、J. H. Fernyhough, A. G. CohnおよびD. C. Hoggは、それ自体が歩行者の軌跡の集まりに関する非パラメータ的な分布の表示である、歩行者広場についての特徴的な行動マップを如何にして習得するかを示したことである。彼らの記述はコンピュータビジョンについてのヨーロッパ学会、1996年、第2巻、475~484ページに掲載されている。彼らのモデルは、本発明の重要な特性のいくつか、特に信号の基本的な構造の簡潔性や正確な回復を欠落している。

【0005】HMMにおける構造習得の研究は、今日まで生成と試験のアルゴリズムに全体的に基づいている。これらのアルゴリズムは、併合または分轄されるべき単一の状態を選択し、次に何らかの利点が得られたかどうかを見るためにそのモデルを再トレーニングすることにより機能する。これらの努力は、失敗を回避するために多様な発見的技法とプライア(prior)を利用しているが、計算の多くは時間を浪費しており、報告される実行時間は数時間から数日の範囲である。Andreas StolckeとStephen Omohundroは、「隠れマルコフモデル帰納法のために併合する最良かつ最初のモデル(Best-first Model Merging for Hidden Markov Model Induction)」(国際コンピュータサイエンス学会技術報告94-003、カリフォルニア大学バークレー校、1994年4月)において、併合アルゴリズムを詳述している。Shiro Ikedaは、「音韻モデルの構造?隠れマルコフモデルのモデル検索(Construction of Phoneme Models - Model Search of Hidden Markov Models)」(インテリジェント信号処理および通信システムに関する国際ワークショップ会報、仙台、1993年10月)において、分轄アルゴリズムを詳述している。

50 【0006】

【発明が解決しようとする課題】隠れマルコフモデルは、信号のモデル化および分類に関して広範に使用されている。ボーム・ウェルチ・アルゴリズムは最尤 (ML) パラメータを効率的に推定するが、ユーザはモデルのグラフィック構造を特定することを義務づけられる。典型的には、ユーザは状態のカウントおよび遷移位相でのいくつかの推測を行うので、各推測を試験することは計算上集約的なものである。

【0007】構造はモデルの選択性および計算速度の主たる決定要素であるので、過程は退屈ではあるが必要なものである。

【0008】

【課題を解決するための手段】本発明においては、動作は隠れマルコフモデルを使って分類され、その精度は、 θ^A の形式でパラメータ推測過程に盛り込まれる構造習得の使用を介して大きく改善されている。エントロピー的プライア (entropic prior) を使用した結果、殆ど情報として役立つ内容がないパラメータを排除または除去することがこれまで不可能であった方法で隠れマルコフモデルを整えることになる。確実なシステムを達成するのに精度を十分改良するだけでなく、処理時間が大幅に削減される。なお、「 θ^A 」は、明細書のコード入力部分において、 θ のべき乗が θ であることを便宜上表す。

【0009】一実施例においては、大きなダイナミックレンジ、すなわち強い確率を備えたパラメータを好む多項式にエントロピー的プライア (entropic prior) が導入されている。情報として役立たないパラメータは消滅させるように、このプライア (prior) を使った反復的な事後最大値 (Maximum A Posteriori) (MAP) 推測はパラメータ値を生成する。HMM トレーニングに適用される場合、その結果得られるモデルは位相的により単純であり優れた一般化を示す。パラメータ消滅の過程を拡大するために、本発明は、低い条件確率がゼロになりうる場合に、全体的に排除する遷移を確立している。最終的には、本発明は不必要な状態をどのように、いつ削除するかを決定する。

【0010】各遷移はモデルを局所的な最大確率から除去して、単純化し、さらにトレーニングを受け易くする。事後確率においては全ての演算は単調であり、したがってトレーニングは直接的に局所的に最適な構造およびパラメータ化に進む。本明細書中に提示した全てのアルゴリズムは、従来のバウム・ウェルチ再推測 (Baum-Welch reestimation) と同じ計算上の複雑性 (例えば、データのサイズで線形であり、遷移の回数で線形である) を有する。

【0011】さらに詳しくは、アルゴリズムは信号における構造の発見のために提供され、粗粒な画像表示から視覚世界の行動に関する標準的な理論を習得するために使われる。この理論は、隠れマルコフモデルの遷移グラフおよび受容領域の形態 (例えば、継続的な出力分布)

を取る。習得アルゴリズムは、構造およびパラメータ推測を素早く同時に行うために、 θ^A の形態でのエントロピー的プライア (entropic prior) を導入し利用する。従来トレーニングされた HMM とは対照的に、この結果得られるモデルは高度に構成されているので人間に理解でき、自動的にフローチャートおよび/または視野における動作パターンなどの特徴的行動のマップに変換することができる。

【0012】本発明においては、システムは中期から長期にかけての周囲の映像から人間の行動のモデルを習得するものである。かかるモデルに必要なのは、そのモデルは視覚データのストリームをコヒーレントな活動に区分すべきであること、異常な行動の検出を可能にすべきであること、および計算上軽量であるべきことである。これらの基準の全ては、粗粒な画像表示の時系列から低エントロピーの有限状態の確率モデルを習得する本発明のアルゴリズムに適合される。さらに、これらのモデルはしばしば、人間に理解できるほど十分に簡潔である。

【0013】さらに、これらのモデルは、フローチャートとしてまたは視野における動作パターンなどの特徴的行動のマップとして読み取り可能な活動の標準的モデルである。

【0014】この結果に関する鍵は、隠れマルコフモデルにおける構造とパラメータの同時習得のための新たな期待値最大化 (expectation-maximization) (EM) アルゴリズムである。

【0015】このアルゴリズムは、状態遷移グラフから情報として役立たない稜 (edge) を切り取ることにより、信号の隠れた構造の低エントロピー表示を見だし、時には状態全体を除去する。このアルゴリズムの基礎は、 θ^A の形式での HMM パラメータ値に関するエントロピー的プライア (entropic prior) と MAP 推定に関する解である。以下で示すように、MAP 推定は、モデルのエントロピーと、そのクロスエントロピーをデータの十分な統計とともに最小化する。データの十分な統計とは、データにおいてそれぞれの異なった種類の事象が発生する時間の分数を意味する。反復的な推測は、情報として役立たないパラメータを消滅させる傾向にあり、これはモデルから切り取ることができる。これにより、簡潔で計算上軽量なモデルが作り出される。実際に、存続する状態はデータの意味のある区分と高度に相関する傾向にあり、一方存続する遷移は信号の力学のうち最も曖昧でないモデルを提供する。

【0016】さらに詳しくは、本発明のシステムは、ビデオカメラおよび/または他のセンサを介して世界を観察し、隠れマルコフモデルの使用により標準的行動の確率モデルを自動的に習得し、そのモデルを使って現在監視している種類の活動を推論して異常な行動を検出する。隠れマルコフモデルの使用は、隠れマルコフモデルのためのパラメータ推測過程に構造習得を組み込み、殆

ど情報として役立たない内容を有するパラメータを隠れマルコフモデルから除去する θ^A エントロピー的トレーナーを備えたモデルのエントロピー的トレーニングにより可能になる。これにより、実時間の確実な異常な行動の検出ができるようになる。一実施例において、このシステムは、画像分析、モデル習得および信号分析の3つの部分で構成される。画像分析において、映像の各フレームは、カメラの前にある対象物の動作を表す数字のベクトルに変形され、一連のかかるベクトルは、映像の各フレームにつき1つ信号を確立する。モデル習得におい

【0017】画像分析は、映像の各フレームにおけるアクティブな画素を識別することによる進行によって始まる。アクティブな画素は世界の移動曲面に属する高い確率を有する。画素を分類するメカニズムは、動作とともに画素値がどのように変化するかについての統計的モデルに基づく。これは、コンピュータビジョンに関する2つの重要な問題であるカメラのノイズと照明の変化に順応する。このシステムは、アクティブな画素をグループ分けしてその累積形状の記述を計算する。良好に機能する2つの記述は、アクティブな画素の (x, y) 位置の第一および第二順位のモーメントと、その $x, y, x+y, x-y$ 位置の個別のヒストグラムである。このシステムは、多種類のセンサから抽出することができる多種類の信号を都合良く使うことができる点に注目することは重要である。例えば、色、音または代替的な動作の特徴を記述する情報は、信号に組み込まれ、信号モデルによって利用することができる。

【0018】信号の習得および分析に使用されるモデルは、上記の隠れマルコフモデルである。本発明のシステムは、隠れマルコフモデルに新たなトレーニング手順を導入している。既存の多くのトレーニング手順と同様に、これは、尤度として知られるモデルと信号との間の適合を最大化する。全ての公知のトレーニング手順とは異なり、これはモデルのサイズの最小化も行う。これにより、良好に一般化され分類される小さく、速いモデルが得られる。トレーニング・アルゴリズムは、オッカムの剃刀が最も偉大な予言力をもって化学的理論を最適な公式に切り落としたように、殆ど情報として役立たない内容を有するパラメータをモデルから除去することにより機能する。これらの情報として役立たないパラメータが残されるとすれば、モデルはトレーニング信号の力学を習得するだけでなく、その偶然の癖や特徴をも習得することになる。これはオーバーフィッティングと呼ばれており、新たな信号が与えられた場合に、習得したモデルがうまく機能しない状態をしばしば引き起こす。

【0019】トリミング(微調整)はオーバーフィッティングを防止するだけでなく、モデルが非常によく構築されているので隠れマルコフモデルを用いて何か新しいことを行うこと、すなわち信号の力学の中で別個の活動を発見することが可能となる。例えば、一実施例において、オフィスの場面を見渡すことから得られる映像信号を習得しながら、このシステムは、電話での会話、ホワイトボードでの作業、コンピュータでの作業および事務椅子でくつろぐことに対応する4種類の異なった動作があることを発見している。習得の後に、このシステムはこれらの活動を検出することができるようになり、この情報は、例えばコンピュータ操作のために照明を暗くする、電話での会話のために雑音を減少させるなどのオフィス環境を補助的に適合させるために使うことができる。さらに、これらのモデルは、病気、動揺、部外者の介入を示唆する異常な行動を検出するという非常によい働きをする。

【0020】

【発明の実施の形態】実施の形態1. ここで図1を参照すると、隠れマルコフモデルを利用して部屋12の中の人物18の異常な行動を検出する、ユニット16に接続されたビデオカメラ14から見た部屋12に関する場面が描写されている。これにより、人物は自分の机20の前の部屋の床に横たわって描写されており、場合によっては、例えば、その人物がてんかんの発作に襲われている場合には明らかに苦痛となる。おそらく椅子22から床へとこの人物が動くという事実はユニット16によって検出され、警報または他の表示がアラーム24を介して行われる。

【0021】標準的でない行動を検出する能力は、前述のエントロピー的プライア(entropic prior)を使ってトレーニングした隠れマルコフモデルを利用した結果である。信号がある時間に亘って標準的特性があるか検査され、標準的特性が確認されると、異常な行動を検出できるのが隠れマルコフモデルの特徴である。

【0022】ここで、図2を参照すると、人物18は椅子22に座っており、また、破線18'で輪郭を図示されるように、ホワイトボード26の前に立っているものとして示されている。人物18はまた、18''で図示されたように、あたかも部屋に入ってきたかのように部屋12に立っているその輪郭を破線で示されてもいる。

【0023】部屋12における全ての活動および動作はカメラ14によって監視されており、異常となる行動よりむしろ、上述のエントロピー的プライア(entropic prior)の利用を通してトレーニングされた隠れマルコフモデルを利用して通常の各種の標準的行動を分類できるように、隠れマルコフモデルのトレーニングによって各種の行動をユニット16で分類できることが理解されよう。この場合には、人物18は自分のモニター28の正面に座っているものとして検出され、矢印30によって

示されるようにその結果生ずる信号は、ユニット32によって印可され、例えば部屋12の照明を暗くする。

【0024】ここで図3を参照すると、場面13は、エントロピー的プライア(entropic prior)でトレーニングされた隠れマルコフモデルの利用を通して、36で図示されるように、システムが標準的行動の確率モデルを最初に習得するユニット16の機能を実行する、点線枠34によって図示された一連のステップにおいて分析される。その後、ステップ38によって図示されるように、現在監視されている特定の行動は分類され、それにより40で図示されるように、異常な行動が検出される。

【0025】ここで図4を参照すると、確率モデルの習得はまず42で図示される画像分析を必要とし、ここで画像はフレーム毎に、場面内の対象または画素の動作を表す数字のベクトルに変えられる。その結果、44で図示されるように、モデル習得が得られる信号を構成する一連のベクトルとなる。46で図示されるように、モデル習得には、隠れマルコフモデルのアルゴリズムの利用を介して確率モデルを得るための信号の分析を含む。確率モデルを得てしまうと、48で図示されるように、信号分析を使って場面の動作を分類して異常を検出する。これは、ユニット42の出力を介して実時間ベースで行われる。

【0026】ここで図5を参照すると、標準的とされるものについての演繹的知識を有することなく標準的行動を素早く判断するシステムの能力は、隠れマルコフモデル52についてのトレーニング手順50を介して可能となる。この手順は、重要でない事象、すなわちその確率が0でも1でもなく、むしろその間の不定位置にある事象を阻止することによって、隠れマルコフモデルのパラメータのトリミングを可能にすることができるよう、エントロピー的プライア(entropic prior)54の利用にまず関連する。したがって、隠れマルコフモデルのトレーニングに使用されるのは、存在する確率が高いか存在する確率が実質的にない事象である。

【0027】エントロピー的プライア(entropic prior)は、56で図示されるように、その事後最大値推定の計算に導入されており、隠れマルコフモデルのトレーニング手順50に直接使用されている。

【0028】演算において、また、隠れマルコフモデルが適切にトレーニングされると、入力信号60は、62に図示されるように、出力が通常でない事象を分類しな

$$Pe(\theta_i) = \theta_i^{\theta_i} = e^{\theta_i \log \theta_i} = e^{-H(\theta_i)}$$

$$Pe(\theta) = \prod_{\theta \in \theta} \theta_i^{\theta_i}$$

【0033】エントロピー的プライア(entropic prior)は0と1で2つの最大値を有する。結果的に、これは曖昧さが低く、ターンオフされる、例えば、0にセットされる多数のパラメータを備えたモデルを好む。図6を参

がらモデルに印可される。

【0029】さらに背景を説明するために、ここで中期から長期にかけての周囲の映像から人間の行動モデルの習得を検討する。かかるモデルに望ましいことは、次のとおりである。視覚データのストリームをコヒーレントな活動度に区分するべきであること、異常な行動の検出を可能にするべきであること、および計算上軽量であるべきこと。粗粒な画像表示の時系列から低エントロピーの有限状態の確率モデルを習得する新しいアルゴリズムにこれらの基準を適合することができると分かる。さらに、これらのモデルは、しばしば人間に理解できるほど簡潔である。これらは、フローチャートとしてあるいは視野における特徴的行動(例えば、動作パターン)のマップとして読み取り可能である行動の標準的モデルである。

【0030】この結果に対する鍵は、隠れマルコフモデルにおける構造およびパラメータの同時習得に関する、新たな期待値最大化(EM)アルゴリズムである。このアルゴリズムは、状態遷移グラフから情報として役立たない部分を切り取ることにより信号の隠れた構造の低エントロピーの表示見だし、時には状態全体を削除することがある。このアルゴリズムの基礎は、パラメータ値に関するエントロピー的プライア(entropic prior)と、事後最大値(MAP)推定に関する解である。MAP推定は、データの十分な統計を用いてモデルのエントロピーとそのクロスエントロピーを最小にする。反復的推測は、情報として役立たないパラメータを消去する傾向にあり、これらはモデルから切り取ることができる。これによって、簡潔で計算上軽量のモデルが作り出される。実際に、存続する状態はデータの意味のある区分と高度に相関する傾向にあり、一方、存続する遷移は信号の力学のほぼ最小限の混乱モデルを提供する。

【0031】エントロピー的推測においては、不確実性を削減しないパラメータは起こりえないと主張される。パラメータは、モデルに何も情報を付加しないことがあり、したがって自由度が消耗する。0または1に近いパラメータは、モデルによってカバーされた信号の分類に強い制約を課すので、情報として役立つ。この洞察力は、パラメータ・エントロピーに対して単純な関係を有する次のようなすっきりした形態で得ることができる。

【0032】

【数1】

(1)

(2)

照すると、図示された曲線は、判断の基礎になる証拠が殆どない場合にシステムが非常に強い確率を好むように、エントロピー的プライア(entropic prior)を使ってシステムを偏向させて確率を割り当てる効果を示す。例

えば、図 6 の曲線の頂点が、2 種類の事象 A と B に関する証拠を与えられた場合、A が B の 2 倍の頻度で発生するように、最良の確率を示す。かかる事象の多くが観察されている場合、推定されるパラメータは、実際に A が起きる十分な確率である $2/3$ に収束する。一方、事象が殆ど観察されていない場合には、システムはデータを信用できず、その推定を事象 A に関してより高い確率の方に歪曲させることを好む。低い位置の太い曲線は、何ら証拠がない場合にはプライア (prior) は 0 または 1 を好むことを示している。

【0034】したがって、 θ^A の使用は、如何なる習得アルゴリズムにもモデルを強制的に単純化させ、パラメータ値を 0 または 1 の方に押しやる。 θ^A はオッカムの剃刀の最適な数学的形態であることが証明できる。モデルとはデータの理論であり、その規則性および構造を記述するものである。 θ^A を使うことにより、一方で単純過ぎてデータにおける基本的情報をすべて捉えることができない理論を有することと、他方で複雑過ぎて回避されるべきデータの偶然的特性を記述する理論を有するこ

$$Pe(\theta | \omega) \propto Pe(\theta) \prod_i \theta_i^{\omega_i} = \prod_i \theta_i^{\omega_i + \omega_i}$$

【0039】このプライア (prior) は明らかに多項式と共役しており、したがって、 ω が証拠となることも分かり、この場合には事後確率は式 (3) と同じ形態を取る。図 6 が示すように、わずかな証拠では、この分布は強い確率の方に偏るが、証拠が増えるにしたがって、 ω について「公平な」確率に収束する。注目すべきは、これがディリクレのプライア (prior) から得られるものと反対の行動であり、これはデータがわずかな場合には弱い確率に偏ることである。

$$\begin{aligned} \theta &= \frac{\partial}{\partial \theta_i} (\log \prod_i \theta_i^{\omega_i + \theta_i} + \lambda \sum_i \theta_i) \\ &= \sum_i \frac{\partial}{\partial \theta} (\omega_i + \theta_i) \log \theta_i + \lambda \sum_i \frac{\partial}{\partial \theta_i} \theta_i \\ &= 1 + \frac{\omega_i}{\theta_i} + \log \theta_i + \lambda \end{aligned}$$

【0043】 θ_i を解析すると、次の式が得られる。

【0044】

$$\theta_i = \frac{-\omega_i}{W_{-1}(-e^{\lambda + 1 + \log \omega_i})}$$

【0045】ここで W_{-1} はランバートの W 関数の二次枝であり、 $W(x) e^{W(x)} = x$ に対する多値解である。式 (5) 及び式 (6) は、エントロピー的 MAP 推定に関する素早い反復の手順を作り出す。 λ を代入して θ を計算し、 θ を標準化し、 θ を代入して λ を計算し、これを繰り返す。これは典型的には 2 ~ 3 回の反復で収

束する。

【0035】簡単に言えば、プライア (prior) は、後に経験によって緩和される強い意見を持つ方がよいと述べており、正確な予測はより多くの信頼性を得て、不正確な予測は習得に関するより診断的情報を提供する。

【0036】上記の θ^A の使用により、図 7 (a)、図 7 (b)、図 7 (c) に関連して以下で明らかのように、何れかの確率モデルからの余分なパラメータの切り取りが可能になる。

10 【0037】さらに詳しくは、図 6 の凸型の曲線は、このプライア (prior) が確率値を嫌い 0、1 の極値を好む様子を示している。2 パラメータの多項式については、最小値は $1/2$ であり、3 パラメータについては $1/3$ であるなどである。これを多項式と組み合わせると、次のように偏向されたエントロピー的プライア (entropic prior) が得られる。ここで、 ω_i は事象タイプ i の偏向である。

【0038】

【数 2】

(3)

【0040】エントロピー的プライア (entropic prior) をすでに導入しているので、ここで確率モデルをトレーニングするのに使われる MAP 推定を生成しなければならない。

【0041】MAP 推定を得るためには、確実に $\sum_i \theta_i = 1$ となるように言語乗数を使って、対数尤度の導関数を 0 に設定する。

【0042】

【数 3】

(4)

(5)

40 【数 4】

(6)

束する。

【0046】式 (3) の対数関数の操作には、エントロピーの点から MAP 推定を理解することを可能にするものもある。

【0047】

【数 5】

50

$$\begin{aligned} \log \prod_i^N \theta_i^{\theta_i + \omega_i} &= \sum_i^N (\theta_i + \omega_i) \log \theta_i \\ &= \sum_i^N (\theta_i \log \theta_i + \omega_i \log \theta_i - \omega_i \log \omega_i + \omega_i \log \omega_i) \end{aligned} \quad (7)$$

$$\begin{aligned} &= \sum_i^N \theta_i \log \theta_i - \sum_i^N \omega_i \log \frac{\omega_i}{\theta_i} + \sum_i^N \omega_i \log \omega_i \\ &= -H(\theta) - D(\omega \| \theta) - H(\omega) \end{aligned} \quad (8)$$

【0048】このように、MAP推定は、パラメータ・エントロピー $H(\theta)$ と、パラメータ θ とデータの十分な統計 ω との間のクロスエントロピー $D(\omega \| \theta)$ との和を最小にする。

【0049】エントロピー的推測は、パラメータ値を漸近に0の方に近づけ、モデル構造が変化するようにして証拠を再分布すること、または数値の丸めのために、パラメータの中には実際に消滅に達するものもある。0

$$\frac{p(\theta \setminus \theta_i)}{p(\theta)} \geq \frac{p(X | \theta)}{p(X | \theta \setminus \theta_i)} \quad (9)$$

【0051】プライア (prior) に代入して、次の対数関数を得る。

$$\log \frac{1}{\theta_i^{\theta_i}} \geq \log p(X | \theta) - \log p(X | \theta \setminus \theta_i) \quad (10)$$

【0053】 θ_i が小さいならば、微分を右項に代入することができる。

$$-\theta_i \log \theta_i \geq \theta_i \frac{\partial \log p(X | \theta)}{\partial \theta_i} \quad (11)$$

【0055】 $-\theta_i$ で割って累乗すると、次の式が得られる。

$$\theta_i \leq \exp \left[-\frac{\partial \log p(X | \theta)}{\partial \theta_i} \right] \quad (12)$$

【0057】したがって、対数尤度の勾配は非常に大きくすることができるので、 θ_i は実際に非常に小さくなくてはならない。しかし、エントロピー的推測は多くのあるいは殆どのパラメータ値をトリミングレンジ内に入れる。

【0058】エントロピー的プライア (entropic prior)、そのMAP推定量およびトリミング演算をすでに

$$\gamma_{ji} = \sum_t \alpha_j(t) P_{ij} p(y_{t+1} | s_t) \beta_i(t+1) \quad (13)$$

【0061】ここで、 α と β は、「隠れマルコフモデルおよび会話認識における選択された応用に関する指導書 (A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition) (IEEE会報、第77巻、第2号、257~286ページ、198

50 はモデルの曖昧さと計算上の複雑さを減少させるので、非常に好ましい。尤度の損失 $p(X | \theta \setminus \theta_i)$ がプライア (prior) $p(\theta \setminus \theta_i)$ における利得によって均衡する時、たとえば下記の数式を検出することにより、確率質量の損失することなくパラメータ θ を明示的に消滅させることができる。

【0050】

【数6】

【0052】

【数7】

【0054】

【数8】

【0056】

【数9】

展開しているので、これらを隠れマルコフモデルのトレーニングに適用することができる。

【0059】HMM遷移確率のエントロピー的推測において、証拠として各遷移についての確率質量を使う

40 【0060】

【数10】

9年)においてRabinerが論述するように、前方・後方分析から得られる。MAP推定量を各 $\omega = \{\gamma_{j,i}\}_i$ に適用することにより、新たな推定 $\{P_{i,j}\}_i = \theta$ を計算する。対照的に、バウム・ウェルチ再推測においては、単に $P_{i,j} = \gamma_{j,i} / \sum_i \gamma_{j,i}$ を設定

する。

【0062】エントロピー的なトレーニングは次のような多くの興味深い特性を有している。すなわち、

(1) より小さな遷移確率は0に近づけられ、その点で遷移全体がモデルから削除されて、計算時間および曖昧さが減退する。

(2) 状態出力分布はより厳密な共分散を有する傾向があり、状態は信号の領域であることをより明確に識別される。

(3) エントロピー的にトレーニングされたHMMは、試験データを提供すると仮定すると、従来のトレーニングされたHMMより高い事後確率を有する傾向にあ

$$P_{ij} \leq \exp \left(-1 - \sum_{t=1}^{T-1} \gamma_j(t) \right)$$

【0065】ここで、 $\gamma_j(t)$ は時間 t における状態 j の確率である。これによりモデルは局所的な最大確率からはじき出されて、さらにトレーニングを可能にする。同様の試験は状態の削除を許諾するが、エントロピー的なトレーニングは、遷移グラフに共通のサブパスを表示するために、いくつかの余分な状態を当然保有する。これにより、モデルのコード付けの長さで計算上の費用が削減される。本明細書中で使用しているように、これらの状態は、殆どどこでもその出力確率は0に近く典型的には計算される必要がないので、データモデル化状態に対してゲーティング(gating)と呼ばれている。

【0066】本発明のシステムの演算の理論をすでに展開しているので、その使用は、図1～5で描写したように、オフィスでの活動のモデルを習得するように適用されてもよい。本明細書中で使用しているように、出力分布および受容領域という言葉は、両方とも状態が同調される信号の領域を意味するので、相互交換可能に使用されている。

【0067】エントロピー的にトレーニングされたHMMは、典型的には単純な身振りの認識作業において優れた性能を提供する。しかし、簡潔な構造的モデルを容易に発見できるために、より長期の行動の構造を習得するのに使用してもよい。例えば、歩行者広場、交差点、工場のフロア、動物のコロニーなどで人がどのように行動するかは、興味ある分野である。オフィスでの活動は、時間の長さが挑戦的な範囲にあるので、特によい試験である。電話を取るといった素早い事象はほんの数秒しかかからず、一方、文書を書くことなどの行為は数時間かかる。この結果は、この構造の多くは軽量で粗粒な視覚追跡を介して発見できることを以下に示している。

【0068】画像の表示に関しては、HMMは、各画像の内容を表示する適度に短い観測ベクトルを必要とする。異なった画像表示は、データ内で異なったコヒーレンシーを強調するモデルにつながることになる。実験は2種類の観測ベクトルであるストライプ表示とプロブ表

り、したがってより正確に分類される。

(4) エントロピー的にトレーニングされたHMMは、初期条件に拘わらず同じ低い混乱度を達成する傾向があり、一方、従来のトレーニングされたHMMの混乱度は、その初期の状態カウンタの関数である。

【0063】トリミングの展開に続いて、HMMにおいては0の近くに残留しているパラメータも、以下の場合および以下の場合にのみ、確率質量の損失なく、削除することができる。

【0064】

【数11】

(1 4)

示を用いて行われた。ストライプ・データは、画像を横切る垂直または水平のストライプにおけるアクティブな画素の平均の位置と平均の範囲で構成されている。プロブ・データは、画像におけるアクティブな画素の単一で最大の接続集合に適合する楕円パラメータで構成されている。

【0069】両方の場合において、アクティブな画素は、背景構造の静的な統計的モデルおよび前景の適応性のあるガウスの色/位置モデルを獲得することによって識別され、これは、すなわち表向き動作によってすでに変化している画素を意味する。画素は尤度比にしたがって前景または背景に分類され、C. Wren, A. Azarbayejani, T. DarrellおよびA. Pentland著「Pfindex: 人体の実時間追跡(Pfindex: Real-time Tracking of the Human Body)」(SPIE会報、第2615巻、1995年)による論文に記述されているように、形態学的な拡大が以前のフレームからのシードを使って前景の画素に接続している。ストライプ・データに関しては、各方向に5～10のストライプが使われていた。観測ベクトルは、各ストライプについて [mean, extent, Δ mean, Δ extent] で構成されている。プロブ・データに関しては、単一の楕円がアクティブな画素に適合された。観測ベクトルは、[mean_x, mean_y, Δ mean_x, Δ mean_y, mass, Δ mass, elongation, eccentricity]で構成されている。

【0070】約30分のデータが(SGI IndyCam)から4Hzで得られた。被験者が部屋と視野から退出したときの空白フレームが自動的に削除されると、ほぼ21分のトレーニングデータが残った。

【0071】期待されるように、プロブ・データに作られたモデルは、主に位置、動作および全体的形状に同調した受容領域を有しており、ストライプ・データに作られたモデルもまた、例えばホワイトボードに書くあるいは電話を取るために腕を伸ばす場合など、体の関節に感知しやすいものであった。両方の場合に、結果は同様に

あった。明細書の以下の部分では、結果は明らかな視覚化に役立つので、プロブ・データに集中する。

【0072】トレーニングに関しては、長さが100から1900フレームに渡るシーケンスが、12、16、20、25および30状態のHMMのエントロピー的トレーニングに使用された。状態は画像にその受容領域とともにタイルを張り付けるように初期化された。遷移確率は隣接するタイルへの動作を好むように初期化され、最初の状態の確率は稜(edge)のない状態については0に設定された。これは図7(a)に描写されているが、そこでは暗い四角はパラメータマトリックスにおける大きな値を示し、明るい四角は0に近い値を示す。初期の受容領域または状態カウントにおける変動は、最終モデルの全体構造または性能において相違がほとんどないことが分かった。トレーニングはMatlabを作動させるSGIR10000で6秒かかった。

【0073】これらの実験の結果は次のとおりである。エントロピー的なトレーニングは、図7(b)に示したように、ほぼ単純化されたパラメータマトリックスを作り出した。これは、フローチャートまたは有限状態の機械の点から特徴的なオフィス活動の高度に読み取り可能で翻訳可能な表示に容易に変換される。図9は、このフローチャートと、その状態とオフィスの区域との間の関係を示している。注目すべきことは、図9の有限状態の機械は、重要でない弧はすでに除去されているのでエントロピー的なトレーニングは基本的な活動をすでに発見していること、およびそれらは選択的にのみ接続されていることである。これは、エントロピー的なトレーニングがない場合の結果である図10から容易に理解できる。いくつかの状態を特に説明することにする。状態5は、データをモデル化せずその他の状態との間のパスを単純化するゲーティング状態である。状態7は、主に延長部に対応し、立ち上がったたり座ったりすることを表す。状態10は、スクリーンを凝視することを表す。状態9は、キーボードを見下ろしたりキーボードから見上げたりすることを表す。

【0074】対照的に、従来のトレーニングは、図7(c)の濃いパラメータマトリックスにおいて見られるような、まばらな遷移グラフで翻訳可能なモデルを作り出さない。従来のトレーニングは、構造化されたモデルを発見できないが、エントロピー的なモデルの徴候はかすかに見える。図10に示された同等な有限状態の機械は、上述のように、全く読み取り不可能である。

【0075】なお、遷移の殆どは切り取られ、多くのオフィスでの活動は、例えばホワイトボードに行ったりホワイトボードから戻ったりするなど対称的な遷移を有するので、遷移の多くは可逆的である。12種類の状態はおそらく準最適のものであるものの、一つの状態はデータのモデル化ではなくゲーティングのために保有された。

【0076】異常な行動を検出するモデルの能力は、エントロピー的にトレーニングされた遷移の重要性を研究するために、いくつかの条件の下で試験された。次の4つのデータ・セットが使われた。すなわち、(a) トレーニング・データ、(b) 提供された試験データ、

(c) 反転して提供された試験データ、(d) 被写体がカップ4杯のエスプレッソを飲んだ後に取ったデータである。これらのデータセットは、原則的に順序、リズムおよび動作のタイミングにおいて異なり、したがって、遷移パラメータの差別的なべき(power)を強調している。次の3つの試験条件があった。すなわち、(1) エントロピー的に推定されたパラメータ、(2) 従来的に推定されたパラメータ、(3) 偶然に平らにされた遷移パラメータである。条件(3)は、遷移または出力パラメータがモデルの選択性に帰因するかどうかを試験する。図8は、データをテストするためによりよく一般化されたエントロピー的なHMMが、異常な行動、例えば後退や神経過敏な行動を区別することに最も成功したことを示している。平らになったモデルの動程は、その選択性が出力パラメータのためだけによるものではほとんどないことを示している。

【0077】要約すると、これまで示してきたのは、画像における変化の極度に粗粒な表示を使って、コンピュータが大量の映像から行動の簡潔な理論をどのように形成できるのかである。30分間のオフィスでの活動の周囲の映像を想定すると、当該システムは、標準的行動の確率モデル、作業活動の読み取り可能なフローチャート、視野における重要な事象および過程のマップを作成する。この結果に対する鍵は、エントロピー的プライア(entropic prior)を利用して、HMMに関する構造およびパラメータ推測を同時に行うアルゴリズムである。期待値最大化アルゴリズムは、モデルのエントロピーとそのクロスエントロピーをデータとともに同時に最小化し、モデルの単純化を通して局所的な確率の最大値を脱することができ、単調であり、数秒で収束する。その結果、システム全体は、現代のワークステーションのコンピュータの能力の4分の1未満を使うだけで、その環境の行動を習得し監視することができる。

【0078】本発明の実施例をいくつかとそれに対するいくつかの変形例を説明してきたが、上記が単に説明的なものにすぎず限定するものではなく、例としてのみ提示されていることが当業者には明白なはずである。多くの変形例やその他の実施例は当業者の範囲内にあり、添付の請求項およびそれと同等な部分によってのみ限定される発明の範囲内に入るものと思量する。

【0079】

【発明の効果】この発明に係る行動検出システムは、以上説明したとおり、入力信号から標準的行動の確率モデルを自動的に習得する行動検出システムにおいて、隠れマルコフモデルを提供する手段と、前記入力信号とエン

トロピー的プライア(entropic prior)に基づいて前記隠れマルコフモデルをトレーニングすることにより、前記確率モデルを生成する手段とを備えたので、標準的行動の確率モデルを自動的に習得し、そのモデルを使って、現在監視下にある活動の種類を推論して異常な行動を検出することができる効果を奏する。

【0080】また、この発明に係る行動検出システムは、以上説明したとおり、前記エントロピー的プライア(entropic prior)が θ^A の形式であり、したがって殆ど情報として役立たない内容を有する前記隠れマルコフモデルからのパラメータを消去するので、現在監視下にある活動の種類を推論して異常な行動を検出することができる効果を奏する。

【0081】また、この発明に係る行動検出システムは、以上説明したとおり、前記トレーニングが、前記隠れマルコフモデルに推定されたパラメータを提供する手段を含み、前記推定されたパラメータを提供する手段が、エントロピー的プライア(entropic prior)を含むので、現在監視下にある活動の種類を推論して異常な行動を検出することができる効果を奏する。

【0082】また、この発明に係る行動検出システムは、以上説明したとおり、前記隠れマルコフモデルをトレーニングする手段が、監視下にある場面对応した映像のフレームを提供するビデオカメラと、前記映像の各フレームを前記ビデオカメラの前の対象物の動作を説明する数字のベクトルに変形させる手段とを含み、一連の前記ベクトルが前記入力信号を備えるので、現在監視下にある活動の種類を推論して異常な行動を検出することができる効果を奏する。

【0083】また、この発明に係る行動検出システムは、以上説明したとおり、前記トレーニング手段が、前記確率モデルについてのパラメータを得るための前記エントロピー的プライア(entropic prior)を含むので、現在監視下にある活動の種類を推論して異常な行動を検出することができる効果を奏する。

【0084】また、この発明に係る行動検出システムは、以上説明したとおり、前記エントロピー的プライア(entropic prior)が θ^A の形式であるので、現在監視下にある活動の種類を推論して異常な行動を検出することができる効果を奏する。

【0085】また、この発明に係る行動検出システムは、以上説明したとおり、標準的行動に関する前記入力信号を分類する前記隠れマルコフモデルを含む手段をさらに含むので、現在監視下にある活動の種類を推論して異常な行動を検出することができる効果を奏する。

【0086】さらに、この発明に係る行動検出システムは、以上説明したとおり、前記分類手段によって確立された標準的行動に基づいて、前記入力信号から異常な行動を検出する、前記分類手段に接続される手段をさらに含むので、現在監視下にある活動の種類を推論して異常

な行動を検出することができる効果を奏する。

【図面の簡単な説明】

【図1】 オフィス備品と、明らかにてんかんの発作にみまわれている人がいる場面の図であり、この場面はビデオカメラで監視されており、その出力は隠れマルコフモデルを用いて人間の異常な行動を判断するのに使われる。

【図2】 ホワイトボードに向かっているか、室内に立っているか、コンピュータ端末に向かって座っている人の位置を図示した図1の場面の図であり、ビデオカメラおよび隠れマルコフモデルは人がどの活動に従事しているかを検出して、人がコンピュータの前で作業を行っていることを検出した場合にはライトを暗くする。

【図3】 標準的行動の確率モデルがまず隠れマルコフモデルの使用を介して習得され、次に現在監視している活動の種類の識別が続き、異常な行動の検出がそれに続き、これらはすべて確率モデルからのものである、図1および図2に図示された場面の図である。

【図4】 画像分析、モデル習得および信号分析を含む行動の決定における一連の事象のブロック図である。

【図5】 形式 θ^A を有するエントロピー的プライア(entropic prior)の導入により、異常な事象を識別するシステムのブロック図であり、このシステムは次に事後最大値推定の計算に使用され、さらに隠れマルコフモデルのトレーニング手順に使用される。

【図6】 観察される証拠に θ^A プライア(prior)を適用した結果である、エントロピーの事後最大値推定量によって行われる確率を誇張して示したグラフである。

【図7】 (a)はトレーニング前の隠れマルコフモデルのパラメータマトリックスの描写である。(b)はエントロピー的推定量を用いたトレーニング後の(a)の隠れマルコフモデルのパラメータマトリックスの描写であり、排除しなければエラーの原因となるであろう余分なパラメータの排除を図示している。(c)はエントロピー的推定量を使用しないトレーニング後の(a)の隠れマルコフモデルのパラメータマトリックスの描写であり、余分なパラメータがどのようにモデルに残されているかを図示している。

【図8】 エントロピー的にトレーニングされた隠れマルコフモデルが、従来的にトレーニングされた隠れマルコフモデルよりも異常信号の検出の作業をうまく行うことを示したチャートである。

【図9】 図7(c)の隠れマルコフモデルの遷移マトリックスから生成されたフローチャートであり、エントロピー的なトレーニングが使用されない場合に、全ての状態またはノードが弧を介して互いに接続されたままであることを示している。

【図10】 図1および図2の場面の人の活動の時間的構造を図示する、エントロピー的にトレーニングされた隠れマルコフモデルから生成されたフローチャートで、

19

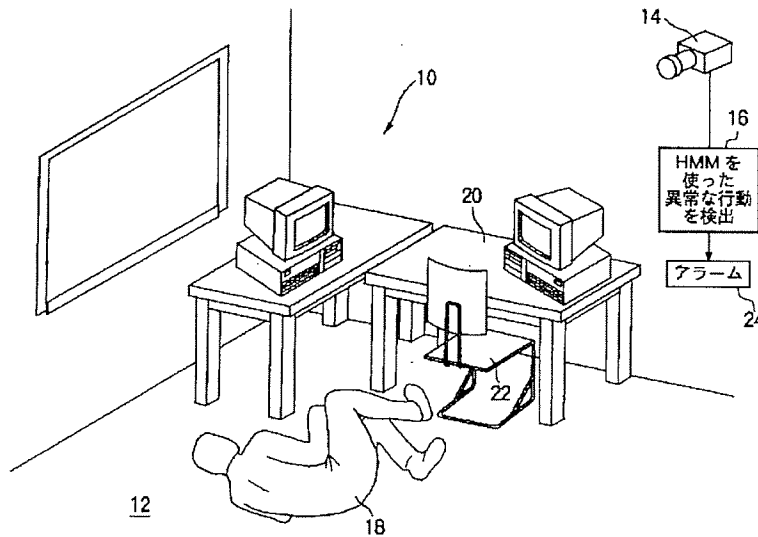
20

元々モデルにあった弧の多くが、各ノードをすでに1ノードおきに接続した弧の全体的なセットから削除されている。

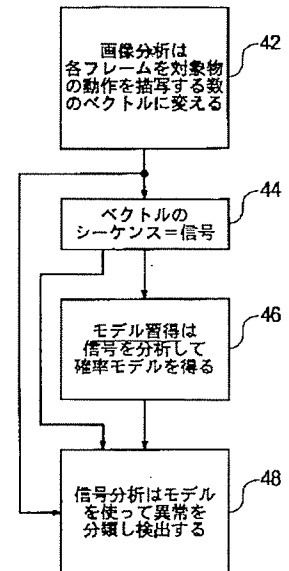
【符号の説明】

12 部屋、14 ビデオカメラ、16 ユニット、18、18'、18'' 人、20 机、22 椅子、24 アラーム、32 ユニット。

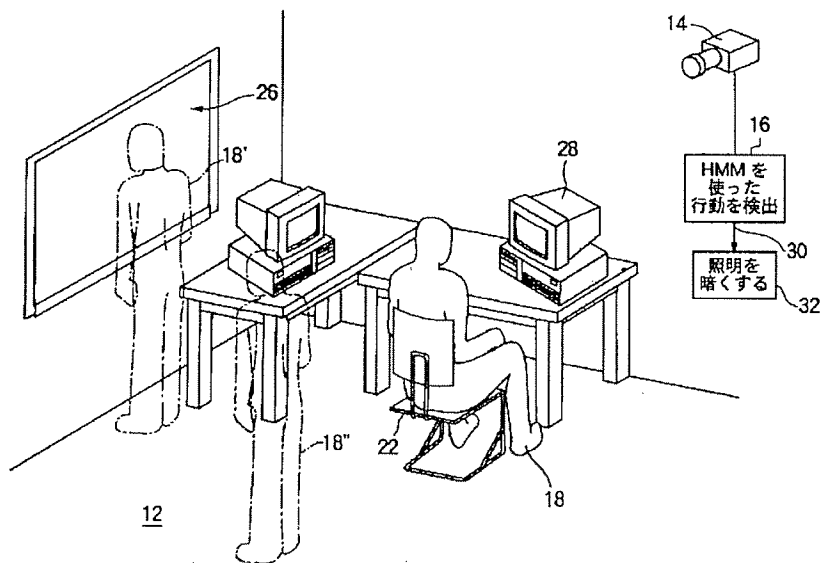
【図1】



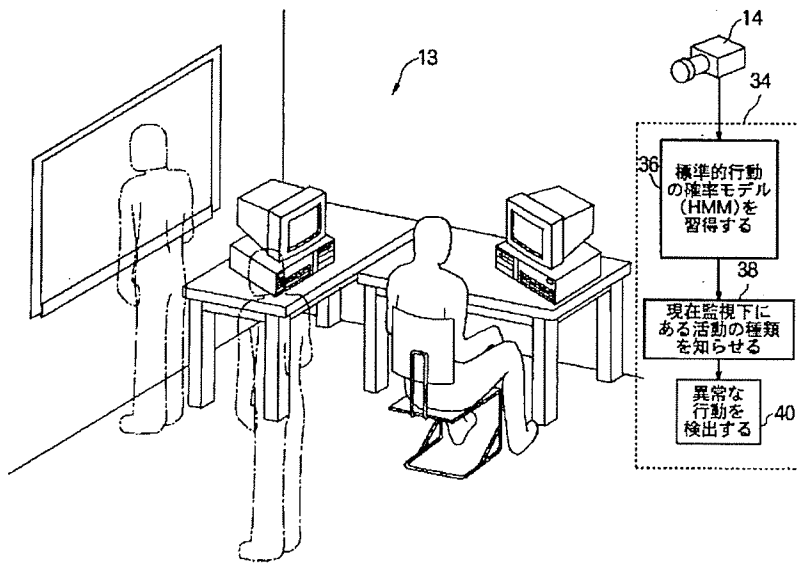
【図4】



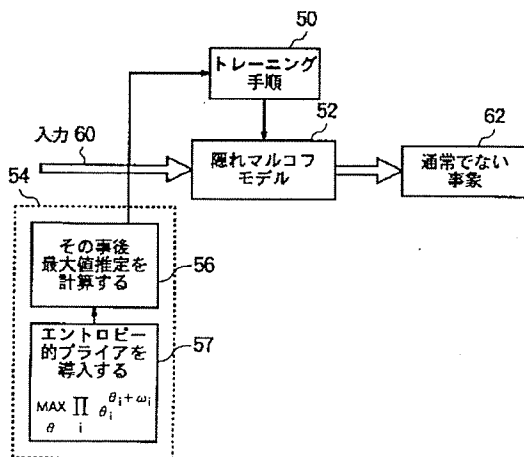
【図2】



【図 3】

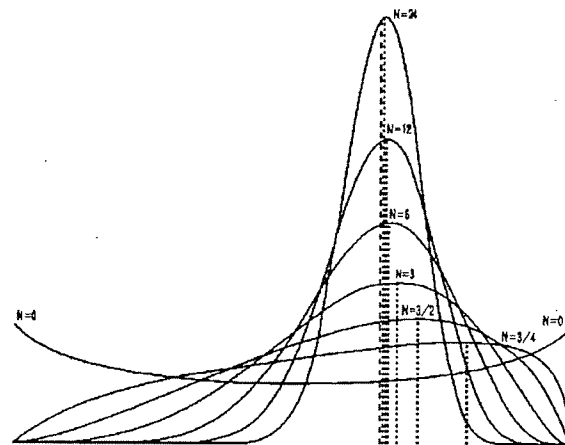


【図 5】

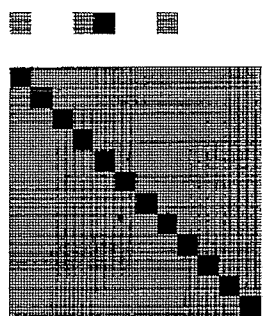
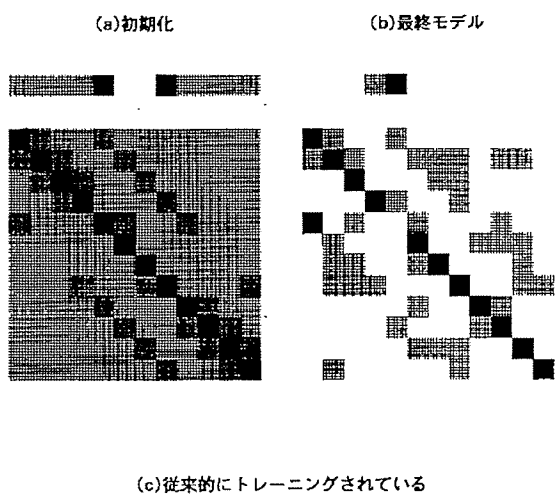


【図 6】

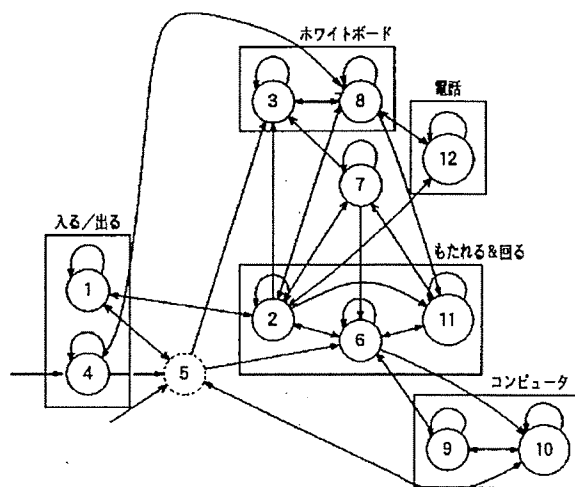
2=1 の比率のエントロピー的分布、N=データポイント総数



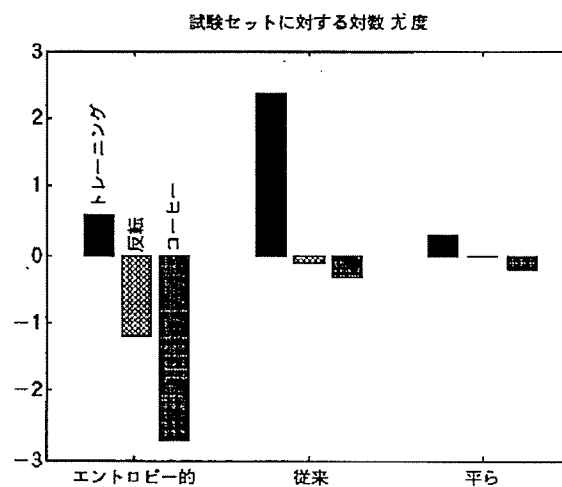
【図 7】



【図 9】



【図 8】



【図 10】

フロントページの続き

(71) 出願人 597067574
201 BROADWAY, CAMBRIDGE,
MASSACHUSETTS
02139, U. S. A.

(72) 発明者 マシュー・イー・ブランド
アメリカ合衆国、マサチューセッツ州、ケンブリッジ、メイプル・アベニュー 28